



JP8125855

Biblio

Page 1

**esp@cenet****INTERSECTION CORRECTING DEVICE FOR TABLE FRAME  
LINE, TABLE RECOGNITION DEVICE, AND OPTICAL  
CHARACTER READER**

Patent Number: JP8125855  
Publication date: 1996-05-17  
Inventor(s): SHINJO HIROSHI; SHIMA YOSHIHIRO; MARUKAWA KATSUMI; KOGA  
MASASHI; NAKAJIMA KAZUKI  
Applicant(s):: HITACHI LTD  
Requested  
Patent: ☐ JP8125855  
Application  
Number: JP19940260857 19941026  
Priority Number  
(s):  
IPC  
Classification: H04N1/40 ; G06K9/20  
EC Classification:  
Equivalents:

**Abstract**

**PURPOSE:** To correct the intersection of the frame line of a table when it is blurred.  
**CONSTITUTION:** The vertical and horizontal rules are detected from the table picture segmented from the inputted picture (202). The outermost frame rule is extracted from among the rules (203). When there is the outer frame rule whose endpoint node is not crossed with the other outer frame rule (206), the outer frame rule is extended (208). When there is an intersection with the other outer frame rule (210), it is made the intersection (212). When there is no intersection with the other outer frame rule (210), the other outer frame rule is extended (214). When there is an intersection with the extended line (216), it is made the intersection (218). Thus, after correcting the intersection of the outer frame rule of the table, the intersection of the inner rule of the table is corrected similarly. Consequently, the intersection of the vanished table frame line can be correctly recovered. Thus, the table and the characters on it can be recognized more accurately and correctly.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-125855

(43) 公開日 平成8年(1996)5月17日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 1/40				
G 0 6 K 9/20	3 4 0 J		H 0 4 N 1/40	C4
			1 0 1 Z	

審査請求 未請求 請求項の数5 O L (全 10 頁)

(21) 出願番号 特願平6-260857

(22) 出願日 平成6年(1994)10月26日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 新庄 広

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 嶋 好博

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 丸川 勝美

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 弁理士 有近 紳志郎

最終頁に続く

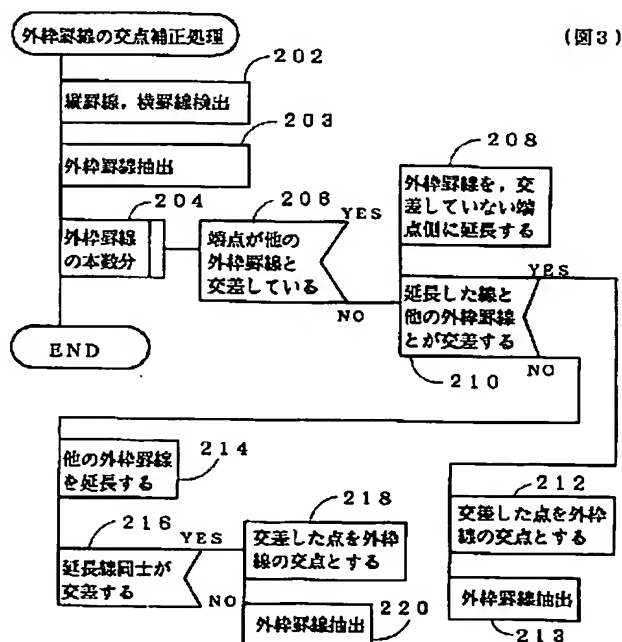
(54) 【発明の名称】 表枠線の交点補正装置および表認識装置および光学文字読取装置

(57) 【要約】

【目的】 表の枠線の交点がかすれている場合に、その交点を補正する。

【構成】 入力された画像中より切り出した表画像から縦罫線と横罫線を検出する(202)。罫線のうち、表の最も外側である外枠罫線を抽出する(203)。端点が他の外枠罫線と交差していない外枠罫線があれば(206)、その外枠罫線を延長して(208)、他の外枠罫線との交点があれば(210)、それを交点とする(212)。他の外枠罫線との交点がない場合は(210)、他の外枠罫線を延長し(214)、その延長線との交点があれば(216)、それを交点とする(218)。このように表の外枠罫線の交点を補正した後、同様に、表の内部罫線の交点を補正する。

【効果】 消えている表枠線の交点を正しく復活させることが出来る。これにより、表や、表の中の文字を正しく認識できる精度が高くなる。



**【特許請求の範囲】**

**【請求項 1】** 画像中の表枠線の交点が消えている場合に、その消えている表枠線の交点を復活させる交点補正装置であって、  
画像中から横罫線と縦罫線とを検出する罫線検出手段と、  
検出した罫線中から表の外枠罫線を抽出する外枠罫線抽出手段と、  
他の外枠罫線との交点が欠落した端点をもつ交点欠落外枠罫線を検出する交点欠落外枠罫線検出手段と、  
交点が欠落した端点側に交点欠落外枠罫線を延長する交点欠落外枠罫線延長手段と、  
前記交点欠落外枠罫線の延長線が他の外枠罫線と交差する場合はその交差した点を外枠罫線同士の交点とする第 1 の外枠線交点復活手段と、  
前記交点欠落外枠罫線の延長線が他の外枠罫線と交差しない場合は他の外枠罫線を延長する他外枠罫線延長手段と、  
前記交点欠落外枠罫線の延長線と前記他の外枠罫線の延長線とが交差する場合はその交差した点を外枠罫線同士の交点とする第 2 の外枠線交点復活手段とを具備することを特徴とする表枠線の交点補正装置。

**【請求項 2】** 請求項 1 に記載した表枠線の交点補正装置において、  
検出した罫線中から表の内部罫線を抽出する内部罫線抽出手段と、  
他の罫線との交点が欠落した端点をもつ交点欠落内部罫線を検出する交点欠落内部罫線検出手段と、  
外枠線の交点を補正し消えていた外枠線を復活させた後に前記交点が欠落した端点側に交点欠落内部罫線を延長する交点欠落内部罫線延長手段と、  
前記交点欠落内部罫線の延長線が他の罫線と交差する場合はその交差した点を罫線同士の交点とする内部枠線交点復活手段とをさらに具備することを特徴とする表枠線の交点補正装置。

**【請求項 3】** 請求項 1 または請求項 2 に記載の表枠線の交点補正装置において、前記罫線検出手段は、両端点ともに他の罫線と交差していない線分を罫線ではなくノイズ成分と判定することを特徴とする表枠線の交点補正装置。

**【請求項 4】** 請求項 1 から請求項 3 のいずれかに記載の表枠線の交点補正装置と、その交点補正装置に表画像を入力する画像入力手段と、前記交点補正装置により補正した交点に基づいて表枠線を認識する枠線認識手段と、その枠線認識手段による認識結果を出力する枠線認識結果出力手段とを具備したことを特徴とする表認識装置。

**【請求項 5】** 請求項 1 から請求項 3 のいずれかに記載の表枠線の交点補正装置と、光学的に読み取った光学読取画像を前記交点補正装置に入力する画像入力手段と、

前記交点補正装置により補正した交点に基づいて表枠線を認識する枠線認識手段と、その枠線認識手段による認識結果に基づいて枠の中の文字行を認識する文字行認識手段と、その文字行認識手段による認識結果に基づいて文字行の中の文字を認識する文字認識手段と、その文字認識手段による認識結果を出力する文字認識結果出力手段とを具備したことを特徴とする光学文字読取装置。

**【発明の詳細な説明】****【0001】**

**【産業上の利用分野】** 本発明は、表枠線の交点補正装置および表認識装置および光学文字読取装置に関する。さらに詳しくは、画像中の表枠線の交点が消えている場合に、その消えている表枠線の交点を復活させる交点補正装置およびその交点補正装置を利用して表を正しく認識する表認識装置および前記交点補正装置を利用して光学読取画像中の文字を正しく読み取る光学文字読取装置に関する。

**【0002】**

**【従来の技術】** 罫線のかすれの補正に関する従来技術としては、特公平 4-18351 公報に記載された技術がある。これは、2本の罫線の端点間の距離と方向の連続性を評価して、距離と方向の連続性の値が基準値以内であれば、2本の罫線を接続して1本の罫線とするものである。表認識における従来技術としては、論文「枠罫線情報を用いた帳票文書の構造認識（成瀬博之その他、電子情報通信学会誌、D-II、Vol. J75-D-I、pp. 1372-1385、1992年）」がある。

**【0003】**

**【発明が解決しようとする課題】** 上記特公平 4-18351 号公報に記載の罫線のかすれの補正方法では、2本の罫線間の距離と方向の連続性を評価している。しかし、表枠線の交点近傍がかすれている場合には、交点を構成する2本の罫線は直交しており、連続性がないため、補正できない問題点がある。また、上記論文「枠罫線情報を用いた帳票文書の構造認識」では、表枠線のかすれについては考慮されておらず、表枠線の交点が消えている場合には、表を誤認識する問題点がある。そこで、本発明の第1の目的は、表枠線の交点が消えている場合に、その交点を正しく復活させる罫線の交点を補正する表枠線の交点補正装置を提供することにある。また、本発明の第2の目的は、前記交点補正装置を利用して、表を正しく認識する表認識装置を提供することにある。また、本発明の第3の目的は、前記交点補正装置を利用して、光学的に読み取った画像中の文字を正しく認識する光学文字読取装置を提供することにある。

**【0004】**

**【課題を解決するための手段】** 第1の観点では、この発明は、画像中の表枠線の交点が消えている場合に、その消えている表枠線の交点を復活させる交点補正装置であ

って、画像中から横野線と縦野線とを検出する野線検出手段と、検出した野線中から表の外枠野線（表の外枠を構成する野線）を抽出する外枠野線抽出手段と、他の外枠野線との交点が欠落した端点をもつ交点欠落外枠野線を検出する交点欠落外枠野線検出手段と、交点が欠落した端点側に交点欠落外枠野線を延長する交点欠落外枠野線延長手段と、前記交点欠落外枠野線の延長線が他の外枠野線と交差する場合はその交差した点を外枠野線同士の交点とする第 1 の外枠線交点復活手段と、前記交点欠落外枠野線の延長線が他の外枠野線と交差しない場合は他の外枠野線を延長する他外枠野線延長手段と、前記交点欠落外枠野線の延長線と前記他の外枠野線の延長線とが交差する場合はその交差した点を外枠野線同士の交点とする第 2 の外枠線交点復活手段とを具備する表枠線の交点補正装置を提供する。

【0005】第 2 の観点では、この発明は、上記構成の表枠線の交点補正装置において、検出した野線中から表の内部野線（外枠野線でない表の野線）を抽出する内部野線抽出手段と、他の野線との交点が欠落した端点をもつ交点欠落内部野線を検出する交点欠落内部野線検出手段と、外枠線の交点を補正し消えていた外枠線を復活させた後に前記交点が欠落した端点側に交点欠落内部野線を延長する交点欠落内部野線延長手段と、前記交点欠落内部野線の延長線が他の野線と交差する場合はその交差した点を野線同士の交点とする内部枠線交点復活手段とをさらに具備する表枠線の交点補正装置を提供する。

【0006】第 3 の観点では、この発明は、上記構成の表枠線の交点補正装置において、前記野線検出手段は、両端点ともに他の野線と交差していない線分を野線ではなくノイズ成分と判定することを特徴とする表枠線の交点補正装置を提供する。

【0007】第 4 の観点では、この発明は、上記構成の表枠線の交点補正装置と、その交点補正装置に表画像を入力する画像入力手段と、前記交点補正装置により補正した交点に基づいて表枠線を認識する枠線認識手段と、その枠線認識手段による認識結果を出力する枠線認識結果出力手段とを具備したことを特徴とする表認識装置を提供する。

【0008】第 5 の観点では、この発明は、上記構成の表枠線の交点補正装置と、光学的に読み取った光学読取画像を前記交点補正装置に入力する画像入力手段と、前記交点補正装置により補正した交点に基づいて表枠線を認識する枠線認識手段と、その枠線認識手段による認識結果に基づいて枠の中の文字行を認識する文字行認識手段と、その文字行認識手段による認識結果に基づいて文字行の中の文字を認識する文字認識手段と、その文字認識手段による認識結果を出力する文字認識結果出力手段とを具備したことを特徴とする光学文字読取装置を提供する。

【0009】

【作用】上記第 1 の観点による表枠線の交点補正装置では、画像中から横野線と縦野線とを検出し（野線検出手段）、検出した野線中から表の外枠野線を抽出し（外枠野線抽出手段）、他の外枠野線との交点が欠落した端点をもつ交点欠落外枠野線を検出し（交点欠落外枠野線検出手段）、交点が欠落した端点側に交点欠落外枠野線を延長し（交点欠落外枠野線延長手段）、前記交点欠落外枠野線の延長線が他の外枠野線と交差する場合はその交差した点を外枠野線同士の交点とし（第 1 の外枠線交点復活手段）、前記交点欠落外枠野線の延長線が他の外枠野線と交差しない場合は他の外枠野線を延長し（他外枠野線延長手段）、前記交点欠落外枠野線の延長線と前記他の外枠野線の延長線とが交差する場合はその交差した点を外枠野線同士の交点とする（第 2 の外枠線交点復活手段）。画像中には、表枠線の線分以外の様々なノイズ成分（例えば文字や表枠線でない線分など）の線分が存在している。このため、画像中の線分を単純に延長して交点を補正すると、ノイズ成分のために誤った交点補正をしてしまう危険がある。例えば、図 2 3 に示すように、野線 4 0 2 を端点 4 0 2 a 側に延長し、文字 “1”

（4 2 6）を端点 4 2 6 a 側に延長し、それらの延長線 4 0 2 b、4 2 6 b の交点 4 2 6 c を表枠線の交点と誤って補正してしまう危険がある。しかし、上記第 1 の観点による表枠線の交点補正装置では、野線検出手段により画像中から横野線と縦野線とを検出することでノイズ成分を除去してしまうため、前記誤った交点補正を回避できる。また、ノイズ成分を除去できても、外枠野線と内部野線とから誤った交点補正をしてしまう危険がある。例えば、図 2 4 に示すように、外枠野線 4 0 2 を端点 4 0 2 a 側に延長し、その延長線 4 0 2 b と内部野線 4 1 8 との交点 4 0 2 c を表枠線の交点とし、また、外枠野線 4 2 0 を端点 4 2 0 a 側に延長し、内部野線 4 0 4 を端点 4 0 4 a 側に延長し、それらの延長線 4 2 0 b、4 0 4 b の交点 4 2 0 c を表枠線の交点とすると、外枠線の交点 4 2 0 d が補正されない危険がある。しかし、上記第 1 の観点による表枠線の交点補正装置では、野線中から外枠野線だけを外枠野線検出により抽出するから、前記誤った交点補正を回避できる。

【0010】上記第 2 の観点による表枠線の交点補正装置では、検出した野線中から表の内部野線を抽出し（内部野線抽出手段）、他の野線との交点が欠落した端点をもつ交点欠落内部野線を検出し（交点欠落内部野線検出手段）、外枠線の交点を補正し消えていた外枠線を復活させた後に前記交点が欠落した端点側に交点欠落内部野線を延長し（交点欠落内部野線延長手段）、前記交点欠落内部野線の延長線が他の野線と交差する場合はその交差した点を野線同士の交点とする（内部枠線交点復活手段）。このように、前記交点欠落内部野線延長手段は、外枠線を復活させた後に内部野線を延長するため、誤った交点補正を回避して、全ての交点を補正できる。

【0011】上記第3の観点による表枠線の交点補正装置では、前記罫線検出手段は、両端点ともに他の罫線と交差していない線分を、罫線ではなく、ノイズ成分と判定する。罫線と平行で且つ罫線と同程度に長い線分が存在する場合、その線分を罫線と間違えて誤った交点補正をしてしまう危険がある。例えば、図25に示すように、線分430を罫線と間違えると、交点430c、410cを誤って補正してしまう危険がある。しかし、両端ともに他の罫線と交差していない線分をノイズ成分と判定して除去するから、前記誤った交点補正を回避できる。

【0012】上記第5の観点による表認識装置では、上記構成の表枠線の交点補正装置を利用して交点を補正し、その補正した交点に基づいて表枠線を認識し（枠線認識手段）、その認識結果を出力する（枠線認識結果出力手段）。従って、正しい表枠線から、表を正しく認識できるようになる。

【0013】上記第7の観点による光学文字読取装置では、上記構成の表枠線の交点補正装置を利用して交点を補正し、その補正した交点に基づいて表枠線を認識し（枠線認識手段）、その認識結果に基づいて枠の中の文字行を認識し（文字行認識手段）、その認識結果に基づいて文字行の中の文字を認識し（文字認識手段）、その認識結果を出力する（文字認識結果出力手段）。従って、正しい表枠線による正しい表認識から、表中の文字を正しく認識できるようになる。

#### 【0014】

【実施例】以下、本発明の一実施例を詳細に説明する。なお、これにより本発明が限定されるものではない。図1は、本発明の一実施例の表枠線の交点補正装置を含む表認識装置を示すブロック図である。この表認識装置20は、イメージスキャナや通信回線を介して画像を入力する画像入力部102と、画像中の表枠線の交点が消えている場合にその消えている表枠線の交点を復活させる交点補正装置10と、その交点補正装置10により補正した交点に基づいて画像中の表枠線の接続関係を抽出しその接続関係に基づいて表枠線を認識する枠線認識部108と、表枠線を認識するためのデータ（例えば表のモデルにおける枠線の接続関係のデータ）を格納している枠線認識辞書110と、前記枠線認識部108による認識結果を出力する枠線認識結果出力部112と、前記画像入力部102から枠線認識結果出力部112までを制御する制御部100とを具備している。なお、枠線認識辞書110の実現例は、前記論文「枠線情報を用いた帳票文書の構造認識」に記載されている。

【0015】図2は、上記表枠線の交点補正装置10による表抽出処理を示すPAD図である。ステップ200では、入力された画像中に含まれる表を輪郭追跡により抽出する。輪郭追跡とは、黒画素と白画素の境界を追跡する処理である。ステップ201では、輪郭追跡して抽

出した表の左上点と右下点とを対角点とする矩形領域を一つの表領域とし、その表領域の画像を切り出す。上記ステップ200、201の処理を、図5～図7に示す具体例により説明する。ステップ200では、図5の画像から、図6に示すような輪郭追跡（600）によって表を抽出する。ステップ201では、図6に示すように、対角点601、602で決まる表領域を切り出す。これにより、図7の画像が得られる。上記表抽出処理によって、隣接する表を分離できる。また、ある程度、ノイズを除去できる。

【0016】図3は、上記表枠線の交点補正装置10による表の外枠罫線の交点補正処理のPAD図である。ステップ202では、表領域から切り出された画像中の縦罫線と横罫線を例えば次の検出規則に従って検出する。

- (1) 罫線は、縦方向または横方向の線分であり、斜方向の線分は罫線ではなくノイズ成分である。
- (2) 罫線は、所定の長さ以上であり、所定の長さより短い線分は罫線ではなくノイズ成分である。
- (3) 罫線は、2つの端点の少なくとも1つが他の罫線と交差しており、2つの端点とも他の罫線と交差していない線分は罫線ではなくノイズ成分である。

【0017】ステップ203では、検出した罫線の中から外枠罫線を例えば次の抽出ルールに従って抽出する。

- (a) ある横罫線上の任意の位置から上を見たときに他の横罫線が常に存在するなら、その横罫線は外枠罫線ではない。他の横罫線が存在しない位置があれば、その横罫線は外枠罫線である。
- (b) ある横罫線上の任意の位置から下を見たときに他の横罫線が常に存在するなら、その横罫線は外枠罫線ではない。他の横罫線が存在しない位置があれば、その横罫線は外枠罫線である。
- (c) ある縦罫線上の任意の位置から左を見たときに他の縦罫線が常に存在するなら、その縦罫線は外枠罫線ではない。他の縦罫線が存在しない位置があれば、その縦罫線は外枠罫線である。
- (d) ある縦罫線上の任意の位置から右を見たときに他の縦罫線が常に存在するなら、その縦罫線は外枠罫線ではない。他の縦罫線が存在しない位置があれば、その縦罫線は外枠罫線である。

【0018】ステップ204では、抽出した外枠罫線の一つに着目してステップ206を実行することを、抽出した外枠罫線の本数だけ繰り返す。ステップ206では、着目している一つの外枠罫線の端点が他の外枠罫線と交差しているかどうかを判定する。交差している場合は、前記ステップ204に戻る。交差していない場合は、交点付近が消えているとみなし、ステップ208に進む。

【0019】ステップ208では、着目している外枠罫線を、他の外枠罫線と交差していない端点側に延長する。ステップ210では、前記ステップ208による延

長線が他の外枠野線と交差するかどうかを判定する。交差する場合には、ステップ 212 に進む。交差しない場合には、ステップ 214 に進む。

【0020】ステップ 212 では、交点を表の外枠線の交点とする。ここで、延長線と交差する他の外枠野線が複数存在するときは、端点に最も近い外枠野線との交点を表の外枠線の交点とする。ステップ 213 では、上記ステップ 203 と同じ外枠野線抽出処理を行う。

【0021】ステップ 214 では、着目している外枠野線以外の外枠野線を延長する。ステップ 216 では、ステップ 208 による延長線と、ステップ 214 による延長線とが交差するか判定する。交差する場合には、ステップ 218 に進む。交差しない場合には、前記ステップ 204 に戻る。ステップ 218 では、交点を表の外枠線の野線の交点とする。ここで、着目している外枠野線の延長線と交差する他の外枠野線の延長線が複数存在するときは延長線が最も短い他の外枠野線との交点を表の外枠線の野線の交点とする。ステップ 220 では、上記ステップ 203 と同じ外枠野線抽出処理を行う。

【0022】上記ステップ 202 からステップ 220 の処理を、図 8～図 17 の具体例により説明する。ステップ 202 では、図 7 の画像から野線を検出する。これにより、図 8 の画像が得られる。ステップ 202 では、図 8 の画像から外枠野線を検出する。これにより、図 9 の画像が得られる。ステップ 206 では、図 10 に示すように、外枠野線 402 の端点 402a が他の外枠野線と交差しているかどうかを判定する。端点 402a は交差していないので、ステップ 208 に進む。ステップ 208 では、延長線 402b を延ばす。ステップ 210 では、延長線 402b と他の外枠野線とが交差するか否かを判定する。延長線 402b は外枠野線 418 と交差するから、ステップ 212 に進む。ステップ 212 では、交点 418a を表の外枠線の交点とする。従って、図 11 に示すようになる。ステップ 213 では、図 11 の画像から外枠野線を検出する。これにより、図 12 の画像が得られる。

【0023】次に、ステップ 206 では、図 13 に示すように、外枠野線 418 の端点 418b が他の外枠野線と交差していないので、ステップ 208 に進む。ステップ 208 では、延長線 418c を延ばす。ステップ 210 では、延長線 418c は外枠野線 410 と交差するから、ステップ 212 に進む。ステップ 212 では、交点 410a を表の外枠線の交点とする。従って、図 13 に示すようになる。

【0024】次に、ステップ 206 では、図 13 に示すように、外枠野線 406 の両方の端点 406a が他の外枠野線 414、420 と交差しているから、次の外枠野線を探してステップ 206 を繰り返す。

【0025】次に、ステップ 206 では、図 14 に示すように、外枠野線 420 の端点 420a が他の外枠野線

と交差していないので、ステップ 208 に進む。ステップ 208 では、延長線 420b を延ばす。ステップ 210 では、延長線 420b は他の外枠野線と交差しないから、ステップ 214 に進む。ステップ 214 では、外枠野線 420 以外の外枠野線を延長する。ステップ 216 では、延長線 420b と他の延長線とが交差するかを判定する。延長線 420b と他の延長線 402d とが交差するので、ステップ 218 に進む。ステップ 218 では、交点 420c を表の外枠線の交点とする。ステップ 220 では、外枠野線を検出する。これにより、図 15 の画像が得られる。

【0026】次に、ステップ 206 では、図 16 に示すように、外枠野線 410 の端点 410a が他の外枠野線と交差していないので、ステップ 208 に進む。ステップ 208 では、延長線 410b を延ばす。ステップ 210 では、延長線 410b は他の外枠野線と交差しないから、ステップ 214 に進む。ステップ 214 では、外枠野線 410 以外の外枠野線を延長する。ステップ 216 では、延長線 410b と他の延長線とが交差するかを判定する。延長線 410b と他の延長線 412b とが交差するので、ステップ 218 に進む。ステップ 218 では、交点 410a を表の外枠線の交点とする。ステップ 220 では、外枠野線を検出する。これにより、図 17 の画像が得られる。

【0027】図 4 は、上記表枠線の交点補正装置 10 による表の内部野線の交点補正処理の PAD 図である。ステップ 300 では、画像から内部野線を検出する。ステップ 301 では、表の内部野線の一つに着目してステップ 302 を実行することを、表の内部野線の本数だけ繰り返す。ステップ 302 では、着目している一つの内部野線の端点 404 が他の野線と交差しているかどうかを判定する。交差している場合は、前記ステップ 300 に戻る。交差していない場合は、交点付近が消えているとみなして、ステップ 304 に進む。ステップ 304 では、着目している内部野線を、他の野線と交差していない端点側に延長する。ステップ 306 では、前記ステップ 304 による延長線が他の野線と交差するかどうかを判定する。交差する場合には、ステップ 308 に進む。交差しない場合には、前記ステップ 300 に戻る。ステップ 308 では、交点を表の枠線の交点とする。ここで、延長線と交差する他の野線が複数存在するときは、端点に最も近い野線との交点を表の枠線の交点とする。

【0028】上記ステップ 300 からステップ 308 の処理を、図 18～図 20 の具体例により説明する。ステップ 300 では、図 18 に示すように外枠線を補正された画像から内部野線 404～408、414～418 を検出する。ステップ 302 では、図 19 に示すように、内部野線 418 の端点 418a が他の野線と交差していないので、ステップ 304 に進む。ステップ 304 では、延長線 418b を延ばす。ステップ 306 では、延



長線 4 1 8 b と他の罫線 4 0 8, 4 1 0 が交差するので、ステップ 3 0 8 に進む。ステップ 3 0 8 では、延長線 4 1 8 b が複数の他の罫線 4 0 8, 4 1 0 と交差するので、端点 4 1 8 a に最も近い他の罫線 4 0 8 を選択し、それとの交点 4 1 8 c を表の枠線の交点とする。以上によって、図 2 0 に示すように、消えていた枠線が復活するので、枠線認識部 1 0 8 は正しく枠線を認識できる。すなわち、上記表認識装置 2 0 によれば、表を正しく認識することが出来る。

【0 0 2 9】なお、上記ステップ 2 1 2, 2 1 8, 3 0 8 において、「交点補正の対象となる端点と罫線もしくは端点同士の間距離が基準値を越える場合には、交点補正をしない」という規則を加えてもよい。

【0 0 3 0】図 2 1 は、本発明の一実施例の表枠線の交点補正装置を含む光学文字読取装置を示すブロック図である。この光学文字読取装置 3 0 は、上述の交点補正装置 1 0 と、その交点補正装置 1 0 にイメージスキャナを介して画像を入力する画像入力部 1 0 2 と、前記交点補正装置 1 0 により補正した交点に基づいて表枠線の接続関係を抽出しその接続関係に基づいて表枠線を認識する枠線認識部 1 0 8 と、表枠線を認識するためのデータ（例えば表のモデルにおける枠線の接続関係のデータ）を格納している枠認識辞書 1 1 0 と、枠の中の文字行を検出する文字行検出部 9 0 2 と、文字行中の文字を文字認識辞書 9 0 6 と照合して認識する文字認識部 9 0 4 と、文字認識結果を出力する文字認識結果出力部 9 0 8 と、前記交点補正装置 1 0 から枠線認識結果出力部 1 1 2 までを制御する制御部 9 0 0 とを具備している。

【0 0 3 1】図 2 2 は、消えていた枠線を復活した表である。文字 4 2 2, 4 2 4, 4 2 6 が表の枠（セル）中にあり、記号 4 2 8 が表中にないことが明確に判る。これに対して、図 5 では、枠線が消えているため、文字 4 2 2, 4 2 4, 4 2 6 および記号 4 2 8 が表中にあるのか否かが明確に判らない。かくして、上記光学文字読取装置 3 0 によれば、表の中の文字を正しく認識することが出来る。

#### 【0 0 3 2】

【発明の効果】本発明の表枠線の交点補正装置によれば、表枠線の交点が消えている場合に、その交点を正しく復活させることが出来る。本発明の表認識装置によれば、表を正しく認識する精度が高くなる。本発明の光学文字読取装置によれば、表の中の文字を正しく認識する精度が高くなる。

#### 【図面の簡単な説明】

【図 1】本発明の一実施例の表枠線の交点補正装置を含む表認識装置を示すブロック図である。

【図 2】本発明の一実施例の表枠線の交点補正装置による表抽出処理の P A D 図である。

【図 3】本発明の一実施例の表枠線の交点補正装置による表の外枠罫線の交点補正処理の P A D 図である。

【図 4】本発明の一実施例の表枠線の交点補正装置による表の内部罫線の交点補正処理の P A D 図である。

【図 5】入力された画像の例示図である。

【図 6】輪郭追跡の説明図である。

【図 7】表領域の切り出し後の画像の例示図である。

【図 8】罫線検出後の画像の例示図である。

【図 9】外枠罫線抽出後の画像の例示図である。

【図 1 0】外枠罫線の交点補正の第 1 例の説明図である。

【図 1 1】外枠罫線の復活の第 1 例の説明図である。

【図 1 2】外枠罫線の交点補正の第 1 例の後の画像の例示図である。

【図 1 3】外枠罫線の交点補正の第 2 例の後の画像の例示図である。

【図 1 4】外枠罫線の交点補正の第 3 例の説明図である。

【図 1 5】外枠罫線の交点補正の第 3 例の後の画像の例示図である。

【図 1 6】外枠罫線の交点補正の第 4 例の説明図である。

【図 1 7】外枠罫線の交点補正の第 4 例の後の画像の例示図である。

【図 1 8】内部罫線抽出後の画像の例示図である。

【図 1 9】内部罫線の交点補正の説明図である。

【図 2 0】消えていた枠線を復活後の画像の例示図である。

【図 2 1】本発明の一実施例の表枠線の交点補正装置を含む光学文字読取装置を示すブロック図である。

【図 2 2】消えていた枠線を復活後の表の画像の例示図である。

【図 2 3】ノイズ成分（文字）による誤った交点補正の例示図である。

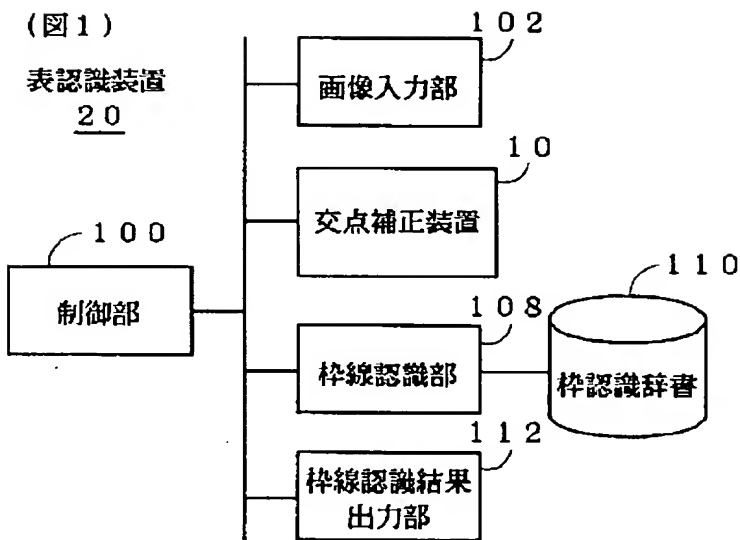
【図 2 4】内部罫線による誤った交点補正の例示図である。

【図 2 5】ノイズ成分（線分）による誤った交点補正の例示図である。

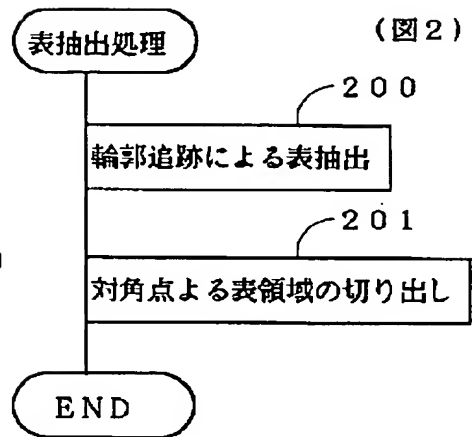
#### 【符号の説明】

1 0	交点補正装置
2 0	表認識装置
3 0	光学文字読取装置
1 0 0, 9 0 0	制御部
1 0 2	画像入力部
1 0 8	枠線認識部
1 1 0	枠認識辞書
1 1 2	枠線認識結果出力部
9 0 2	文字行検出部
9 0 6	文字認識辞書
9 0 4	文字認識部
9 0 8	文字認識結果出力部

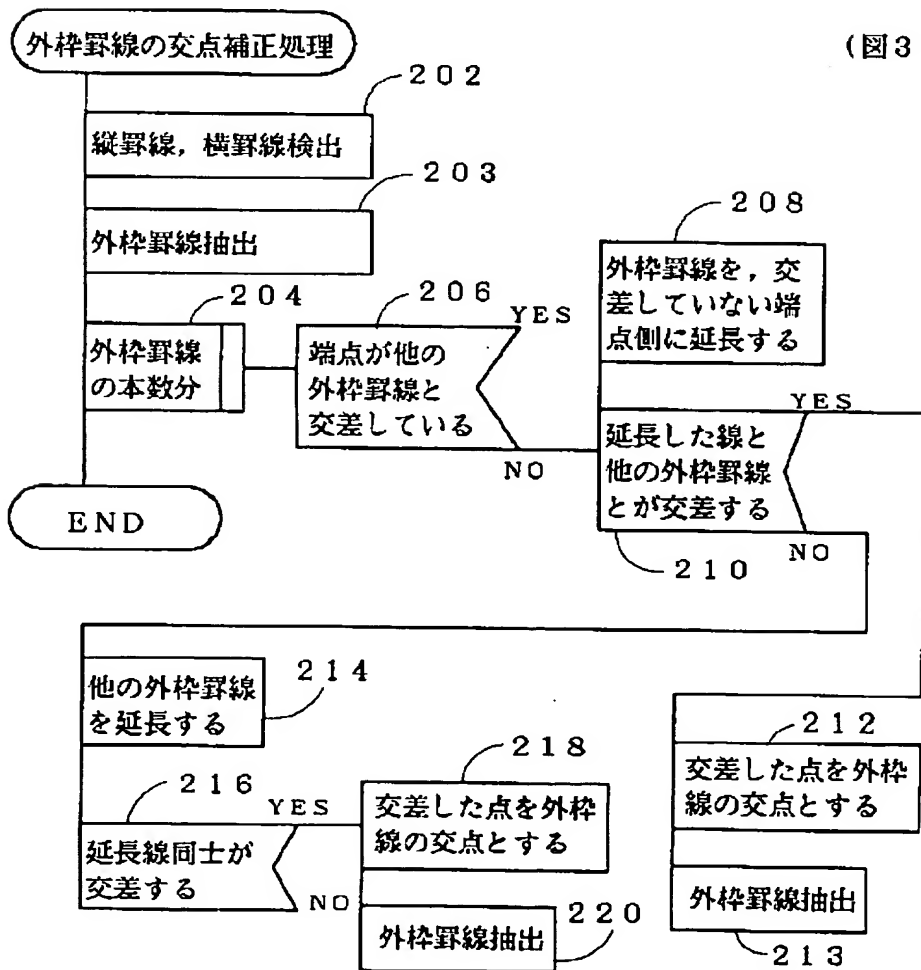
【図1】



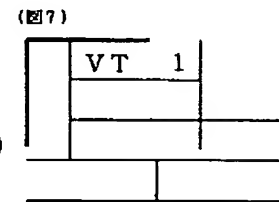
【図2】



【図3】

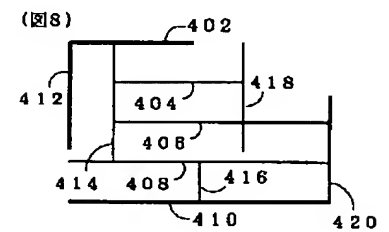


【図7】

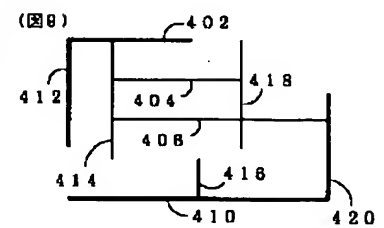


(図3)

【図8】

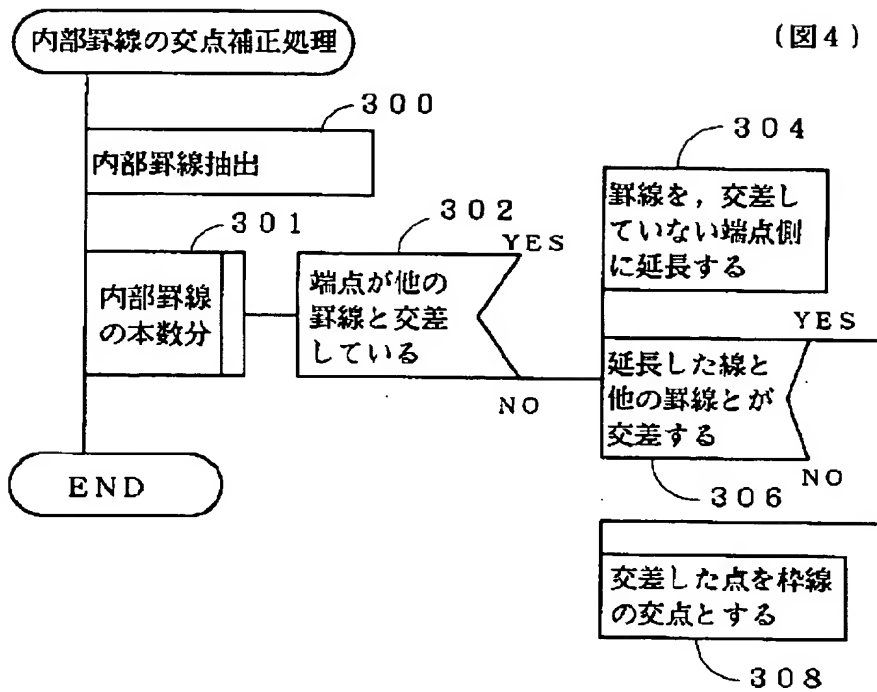


【図9】

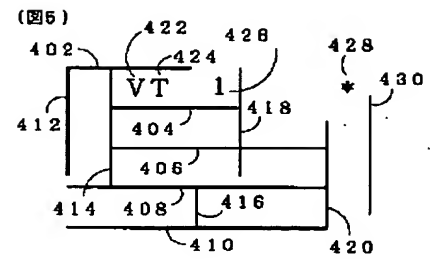




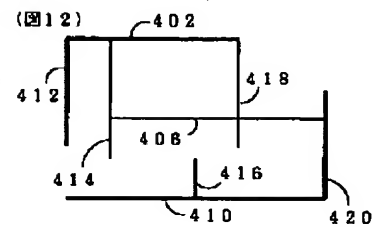
【図 4】



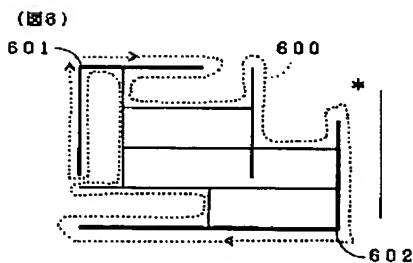
【図 5】



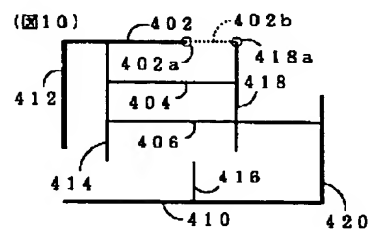
【図 12】



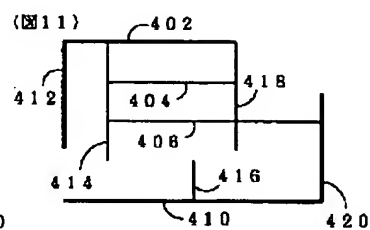
【図 6】



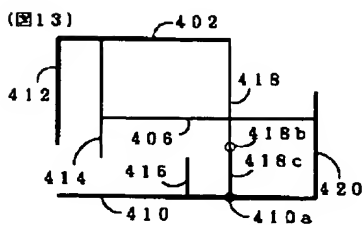
【図 10】



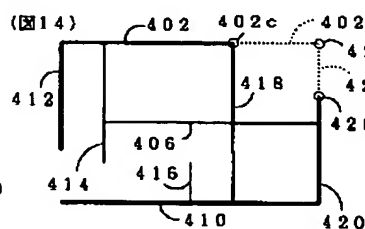
【図 11】



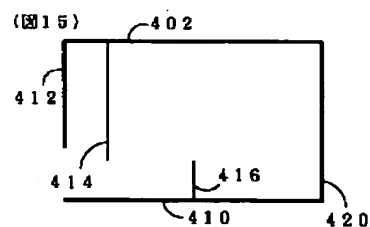
【図 13】



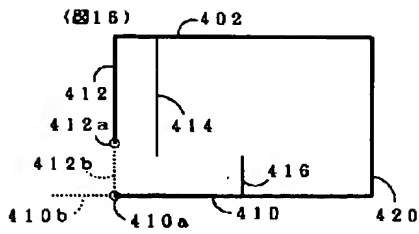
【図 14】



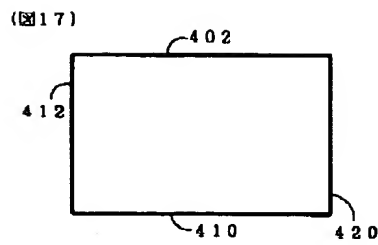
【図 15】



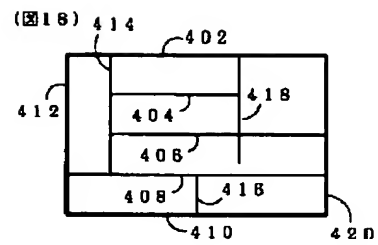
【図16】



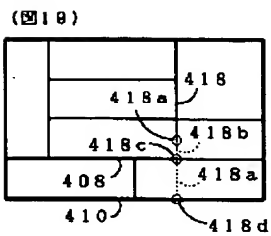
【図17】



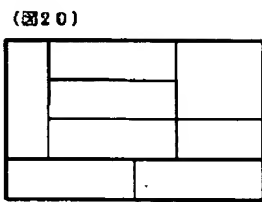
【図18】



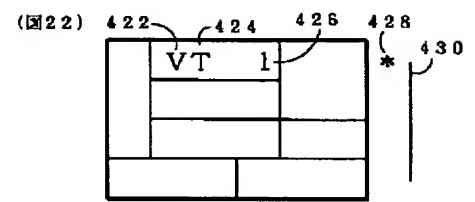
【図19】



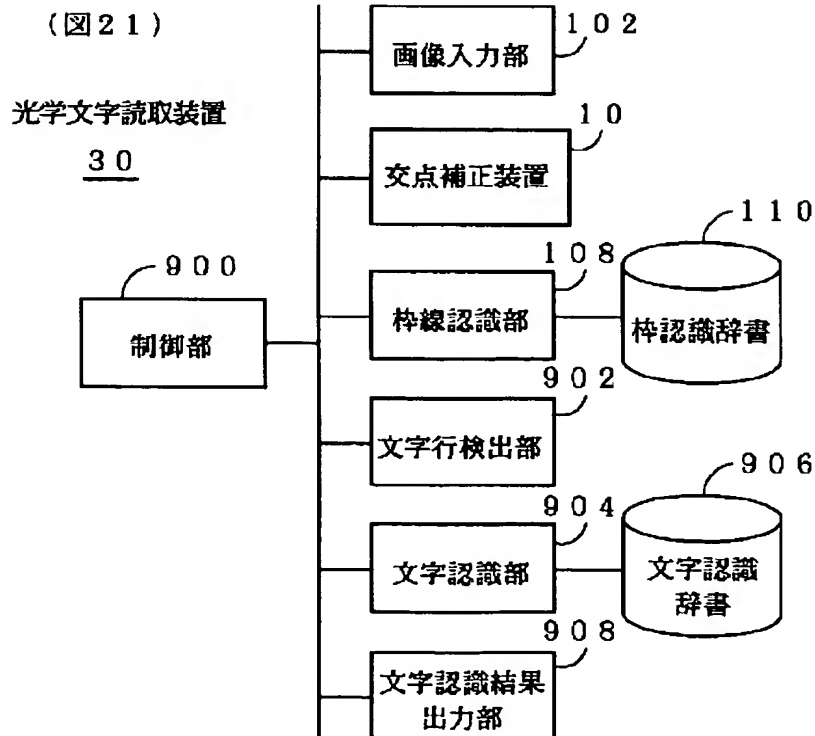
【図20】



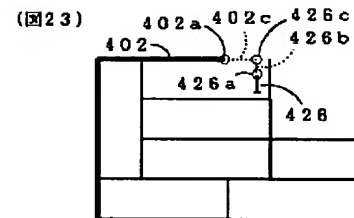
【図22】



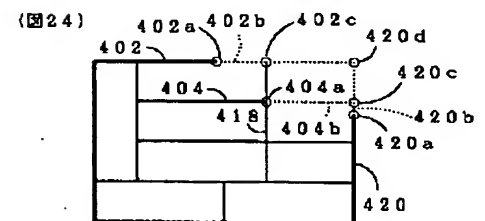
【図21】



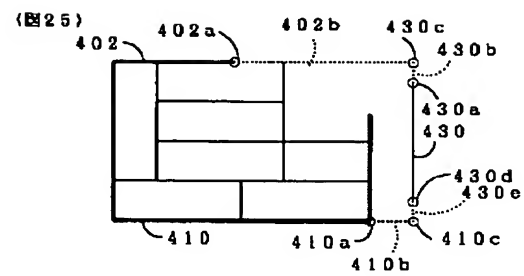
【図23】



【図24】



【図 25】



フロントページの続き

(72)発明者 古賀 昌史  
東京都国分寺市東恋ヶ窪 1 丁目 280 番地  
株式会社日立製作所中央研究所内

(72)発明者 中島 和樹  
東京都国分寺市東恋ヶ窪 1 丁目 280 番地  
株式会社日立製作所中央研究所内